

4.1. METADATOS.

Los metadatos han estado con nosotros desde que el primer bibliotecario hizo una lista de los documentos que tenía en un estante, en un pergamino escrito a mano. El término "meta" viene de una palabra griega que significa "junto a, con, después, siguiente". El uso más reciente del latín y el inglés emplearía "meta" para indicar algo trascendental, o sobrenatural. Los metadatos, pues, pueden ser definidos como datos sobre otros datos. Es el término usado en la era de Internet para la información que los bibliotecarios tradicionalmente habían puesto en los catálogos, y más comúnmente se refiere a información descriptiva sobre recursos de la Web. Un registro de metadatos consiste en un conjunto de atributos, o elementos necesarios para describir la fuente en cuestión. Por ejemplo, un sistema de metadatos común entre los bibliotecarios, el catálogo de biblioteca, contiene un conjunto de registros de metadatos con elementos que describen un libro u otra publicación en una biblioteca: autor, título, fecha de creación o publicación, materia, y la signatura topográfica especificando la localización de la publicación en el estante.

La relación entre un registro de metadatos y el recurso al que describe puede darse de una de estas dos formas¹:

- Los elementos pueden estar en un registro separado del documento, como en el caso del registro de un catálogo de bibliotecas; o,
- Los metadatos pueden estar incluidos, incrustados, en el propio recurso.

El concepto de metadatos antecede a Internet y a la Web pero, como se puede suponer, el interés mundial por las normas y prácticas de metadatos ha estallado con el crecimiento de la publicación electrónica y las bibliotecas digitales, y la concurrente "sobrecarga de información" que resulta de las grandes cantidades de datos digitales disponibles en línea.

Actualmente hay un gran interés, y constituye una parte fundamental del desarrollo del Web semántico, por adoptar a gran escala estándares y

¹ Dublin core Metadata Initiative. Guía de uso del Dublin Core. Disponible en: <http://es.dublincore.org/documents/usageguide/index.shtml> (23/06/04).

prácticas descriptivas para los recursos electrónicos porque ello contribuirá a mejorar la recuperación de recursos relevantes en cualquier contexto.

Como señalan Weibel y Lagoze [Weibel,97], dos líderes en el campo del desarrollo de metadatos:

"La asociación de metadatos descriptivos normalizados a los objetos de la red tiene el potencial para mejorar sustancialmente las capacidades de localización/recuperación, facilitando búsquedas basadas en campos (p. ej. autor, título), permitiendo la indización de objetos no textuales, y facilitando el acceso al contenido sustituido/referenciado que es distinto del acceso al contenido del propio recurso".

Hay varias iniciativas de estandarización de metadatos dependiendo del dominio de aplicación, ya que un sólo estándar no puede recoger todos los requisitos específicos. Por ejemplo, los primeros campos en los que se empezó a investigar el uso de metadatos fueron en humanidades y lingüística a través de TEI², biología³, y en las ciencias geoespaciales⁴. A nosotros nos interesa conocer las iniciativas que tratan de estandarizar metadatos que permitan describir y recuperar recursos educativos. Hay varias iniciativas actualmente en marcha, de las que vamos a describir brevemente las siguientes:

4.1.1. IEEE/LTSC LOM e IMS Global Learning Meta-Data.

Se trata del estándar por excelencia para Metadatos de Objetos Educativos, y está patrocinado por el Comité de Estandarización de Tecnologías Educativas del IEEE. Alrededor de este estándar es habitual encontrarse con confusiones sobre el nombre y sus autores. Veamos cómo surgió:

En 1997 por parte del consorcio EDUCOM que agrupaba varias instituciones universitarias americanas, se puso en marcha el proyecto IMS, cuyo principal objetivo era el de desarrollar estándares de especificación de metadatos de materiales educativos. Por otra parte, en el mismo año el grupo P.1484 del IEEE (llamado ahora IEEE Learning Technology Standards

² Text Encoding Initiative, <http://www.tei-c.org/index.html>

³ National Biological Information Infrastructure.Metadata Home Page. Disponible en: <http://www.nbi.gov/datainfo/metadata/> (23/06/04).

⁴ Federal Geographic Data Committee. Metadata Home Page. Disponible en: <http://www.fgdc.gov/metadata/metadata.html> (15/11/04)

Committee) se propuso el mismo objetivo: estudiar los metadatos adecuados para el marcaje de materiales destinados a la educación.

A su vez, ARIADNE que es un proyecto de investigación europeo que incluía una parte importante de definición de metadatos, e IMS empezaron a colaborar, y fruto de este trabajo conjunto, sometieron a consideración del grupo del IEEE una propuesta que se constituyó después en la base del borrador del IEEE LOM (Learning Object Metadata). En Junio de 2002, LOM fue aprobado como estándar IEEE con la referencia 1484.12.1⁵.

IEEE LOM define y especifica un esquema de metadatos que permite múltiples implementaciones, los atributos, sus definiciones, una estructura jerárquica que los relaciona entre ellos, y por tanto los aspectos teóricos del esquema, pero no incluye información acerca de cómo representar estos metadatos o con qué mecanismos se puede transmitir y procesar esta metainformación. Por ello, en el seno del proyecto IMS se han especificado modelos para el uso de los metadatos y una representación en XML que proporcionan los mecanismos para ser usados en la práctica.

Así, tanto IEEE Learning Object Meta-Data como IMS Learning Resource Meta-Data hacen referencia a lo mismo: IEEE LOM es el estándar promovido por el IEEE que contiene la especificación de los atributos, nombres, definiciones, tipos de datos y longitudes, mientras que IMS Learning Resource Meta-Data incluye una referencia al anterior estándar y además las guías y herramientas necesarias para su uso [Peig,04].

El principal objetivo de este estándar es facilitar la búsqueda, la evaluación, la adquisición y el uso de recursos educativos, tanto por parte de los instructores como de los alumnos. Igualmente se pretende facilitar el intercambio de los "objetos educativos"⁶, permitiendo el desarrollo de catálogos e inventarios y teniendo en cuenta la diversidad cultural de los entornos donde estos recursos y los metadatos asociados pueden ser usados.

Así pues, los estándares del LOM se centran en el conjunto mínimo de propiedades que permiten que los objetos educativos sean gestionados, ubicados y evaluados [Cámara,02].

⁵ IEEE LOM Standard. Disponible en: <http://ltsc.ieee.org/wg12/par1484-12-1.html> (23/06/04).

⁶ Un objeto educativo se puede definir como una entidad, digital o no digital que puede ser usada, reutilizada o referenciada durante cualquier actividad de aprendizaje basada en la tecnología.

Los metadatos LOM se estructuran en una jerarquía en árbol. El nodo raíz corresponde al documento que se está describiendo y suele recibir el nombre de "lom". En el siguiente nivel encontramos sub-elementos, que pueden contener a su vez otros sub-elementos. A los elementos terminales se les llama hojas y a los intermedios, ramas. Para cada elemento en la jerarquía se especifica la definición, el tipo de datos, los valores permitidos y si se permite multiplicidad o no.

El Esquema de Base de LOM se compone de 9 categorías y 47 elementos:

Categorías	Elementos
1. General	Identificador, Título, Entrada de catálogo, Lengua, Descripción, Descriptor, Cobertura, Estructura, Nivel de agregación
2. Ciclo de vida	Versión, Estatus, Otros colaboradores
3. Meta-metainformación	Identificador, Entrada de catálogo, Otros colaboradores, Esquema de metadatos, Lengua
4. Técnica	Formato, Tamaño, Ubicación, Requisitos, Comentarios sobre la instalación, Otros requisitos para plataformas, Duración
5. Uso educativo	Tipo de interactividad, Tipo de recurso de aprendizaje, Nivel de interactividad, Densidad semántica, Usuario principal, Contexto [Nivel educativo], Edad, Dificultad, Tiempo previsto de aprendizaje, Descripción, Lengua
6. Derechos	Coste, Copyright y otras restricciones, Descripción
7. Relación [con otros recursos]	Tipo [naturaleza de la relación con el recurso principal], Recurso [recurso principal al que se refiere esta relación],
8. Observaciones	Persona, Fecha, Descripción
9. Clasificación	Finalidad, Nivel táxon (taxonómico), Descripción, Descriptor

Tabla 2. Esquema de Base de LOM

En la *Figura 5* se muestran los sub-elementos que pertenecen al primer elemento, "general".

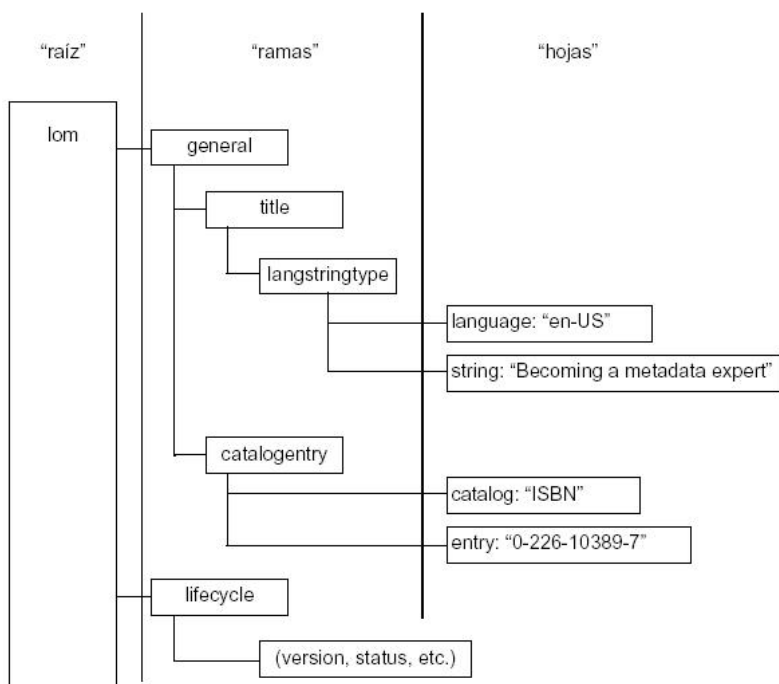


Figura 5. Estructura jerárquica de los metadatos IEEE LOM.

4.1.2. Dublin Core Metadata Initiative.

DCMI (Dublin Core Metadata Initiative)⁷ es una organización internacional y virtual albergada por OCLC (Dublin Ohio, EE.UU) que se financia a través de proyectos y subvenciones y en la que el trabajo se realiza por voluntarios de todo el mundo.

Constituye un mecanismo básico de descripción que puede usarse en todos los dominios, para todo tipo de recursos, sencillo pero potente, que puede extenderse fácilmente y puede trabajar conjuntamente con otras soluciones específicas [Dekkers,03]. Sus objetivos son:

- Desarrollar estándares de metadatos para la recuperación de información en Internet entre diversos dominios informativos.
- Definir marcos de trabajo para la interoperabilidad entre conjuntos de metadatos.
- Facilitar conjuntos de metadatos específicos para una comunidad o disciplina, acordes a los apartados 1 y 2.

Los principales temas de trabajo se concretan en:

⁷ DCMI. Disponible en: <http://dublincore.org/>

- Expresar metadatos Dublin Core Metadata en XML, RDF/XML, etc.
- Extensión y uso del DC a dominios específicos de información (Perfiles de aplicación)
- El Registro DCMI (diccionario)
- Citación, Agentes, Colecciones
- Soporte para el desarrollo de herramientas

De forma general, los metadatos Dublin Core constan de:

- 15 elementos básicos:

Título, Creador, Colaborador, Editor, Fecha de publicación, Identificador (URI), Materia, Descripción, Cobertura, Tipo de recurso, Idioma, Formato, Fuente, Relación con otros documentos, Derechos.

- calificadores para detalles adicionales:

Matizaciones: p. ej. Date.Created (fecha de creación), Relation.IsPartOf (relacionado por ser parte de...).

Vocabularios controlados (p. ej. Encabezamientos de materia normalizados, clasificaciones, nombres geográficos) .

- Tipos de datos (formatos de fecha/tiempo, tipos MIME).

Dublin Core puede, por tanto, verse como un "pequeño lenguaje para realizar una clase particular de declaraciones sobre recursos". En este lenguaje, hay dos clases de términos: elementos (nombres) y calificadores (adjetivos), que pueden ordenarse en un patrón simple de una sentencia o declaración. Los sujetos, implícitos o sobreentendidos, de este lenguaje, son los propios recursos⁸.

En el seno de DCMI se han establecido grupos de trabajo que trabajan en sectores concretos, y de ellos el que nos interesa es el Educación, que está dedicado al estudio de la aplicación de metadatos Dublin Core a recursos educativos y que tiene como objetivo establecer puntos de encuentro con el IEEE LOM. Ambas iniciativas han decidido colaborar en el desarrollo de especificaciones de metadatos para la Web, de forma que:

- Se evite el solape semántico entre ambas propuestas de estándares.

⁸DCMI Principios Gramaticales. Disponible en:
<http://es.dublincore.org/usage/documents/principles/> (23/06/04)

- Se promueva la independencia de cualquier tecnología o sintaxis para expresar metadatos (Sin embargo, piensan usar diversas tecnologías: HTML, XML, RDF).
- Los metadatos han de ser interoperables y reutilizables.

4.1.3. Ejemplo de aplicación de metadatos: SCORM.

La iniciativa ADL⁹ (Advanced Distributed Learning, 1997), es un programa del Departamento de Defensa de los Estados Unidos y de la Oficina de Ciencia y Tecnología de la Casa Blanca para desarrollar, a través del ADL Co-Lab, los principios y guías de trabajo necesarias para el desarrollo y la implementación eficiente, efectiva y en gran escala, de formación sobre nuevas tecnologías Web. ADL Co-Lab no desarrolla especificaciones propias sino que integra y complementa las ya existentes (como IMS del IEEE/LTSC), facilitando, a través de "laboratorios" compartidos con toda la industria, la comprobación y validación de estos estándares emergentes y proponiendo marcos de trabajos comunes que los integre. Todo ello lo han recogido bajo un modelo que denomina SCORM (Sharable Content Object Reference Model.)

Tras un intensa colaboración con IEEE, AICC e IMS, ADL publicó en Enero de 2000 la versión 1.0 del modelo ADL SCORM, que, en realidad, es una selección, compilada y completada de las especificaciones más críticas, urgentes y avanzadas ya en IEE/LTSC (por ejemplo, el interface CMI, que regula la interacción entre cursos y plataforma, originario de AICC y en proceso de estandarización en grupo IEEE/LTSC P1484.11 o los metadatos IMS, en proceso, entonces, de estandarización en IEEE/LTSC P1484.12 como ya hemos comentado). ADL SCORM 1.1 marca un hito significativo en el camino hacia los estándares, que debía servir como base para el desarrollo a gran escala de utilidades y contenidos compatibles, abriendo así, en la práctica, su fase de implementación industrial y uso comercial.

A lo largo de todo el 2000, ADL CO-Lab promovió varias "Plug-Fest" o convocatorias abiertas a industria, universidad, defensa y grupos de estandarización, tanto para evaluar la viabilidad del modelo inicial, como para permitir a todos los interesados sincronizar la evolución y convergencia

⁹ ADL, Advanced Distributed Learning. Disponible en: <http://www.adlnet.org/> (24/06/04).

de todas las herramientas de creación, plataformas formativas y cursos, bajo estos estándares abiertos (fase de prueba o "beta-testing").

Como consecuencia de dicho análisis, en 2001 apareció la versión 1.1 que normalizaba y especificaba, en la práctica, las necesidades más interesantes y críticas para la interoperabilidad de los sistemas de formación virtual, para lo que permitía definir:

- Todos los datos (alumnos, organización, actividad formativa, resultados de aprendizaje, datos de evaluación, etc.) que debe recoger una plataforma formativa ADL SCORM 1.1 compatible.
- Los datos que puede generar el alumno en su navegación y aprendizaje por un curso ADL SCORM 1.1 compatible (teóricos, prácticos, evaluaciones, etc.) y cómo éstos deben enviarse a cualquier plataforma ADL SCORM 1.1 compatible.
- Cómo debe construirse un curso ADL SCORM 1.1 compatible, su estructura, sus posibles itinerarios pedagógicos, restricciones y exigencias, y como calificar con metadatos los objetos educativos.
- Cómo se debe importar/exportar un curso ADL SCORM 1.1 compatible entre cualquier utilidad de creación de contenidos y una plataforma ADL SCORM 1.1 .

Posteriormente han aparecido la Versión 1.2 (que permitía la adición de contenido empaquetado a los perfiles de aplicación, convergiendo con la especificación de paquetes de contenido de IMS), y la última versión, considerada por IMS como estable, denominada SCORM 2004 (Enero 2004) que introduce la secuenciación (como se juntan y secuencian diferentes objetos de aprendizaje para crear unidades de instrucción de mayor nivel) y la navegación.

En resumen SCORM es un programa de estandarización que pretende recoger las premisas y los patrones de trabajo necesarios para desarrollar, a gran escala, la formación a través de las TIC (pasa por ser el primer modelo estándar aplicable a la formación virtual), usando el lenguaje XML para resolver los problemas básicos de carga de un curso en un LMS (Learning Management System) de otro fabricante, estableciendo un modo de desarrollar, empaquetar y gestionar la distribución de unidades formativas digitales con las siguientes características:

- Reusable: modificable por diferentes herramientas

- Accesible: puede ser publicado y encontrado por diferentes entidades y sistemas.
- Interoperable: capaz de funcionar en diferentes sistemas servidor y cliente.
- Duradero (persistente): no requiere modificaciones significativas para adaptarlo a un nuevo sistema.

De lo tratado hasta aquí podemos deducir que no vamos por mal camino en nuestra misión de enriquecer el contenido semántico de los WebQuest. Necesitamos recurrir a los metadatos para lo que existen dos posibilidades, o IMS/LOM o DUBLIN CORE, aunque la primera parece la más aconsejable en nuestro caso dado que se dedica exclusivamente a objetos educativos. Además, hemos visto que ya hay concreciones operativas en el campo de la distribución virtual de contenidos educativos (que es en el que podríamos encuadrar nuestra misión), de lo cual hemos visto como ejemplo SCORM.

SCORM es una iniciativa americana que tiene su doble en Europa denominado ARIADNE¹⁰, y que se dedica a la definición de metadatos e intercambio de recursos educativos; tiene suscrito un acuerdo de colaboración con SCORM, y con IEE LOM (o IMS) con fines de estandarización, como ya se ha expuesto.

Como se mostró en el epígrafe anterior, tanto IEEE LOM (IMS) como DUBLIN CORE promueven el uso del lenguaje XML (eXtensible Markup Language) y RDF (Resource Description Framework) para representar metadatos. Si queremos usar metadatos conviene, entonces, estudiar en qué consiste el lenguaje denominado XML.

[Weibel ,97] Weibel, S., Kunze, J., & Lagoze, C. (1997, February 9) Dublin Core Metadata for simple resource description.

[Peig,04] Peig, Enrique. Interoperabilidad de Metadatos en Sistemas Distribuidos. Tesis Doctoral UPF. Disponible en: <http://www.tdx.cbuc.es/TDX-0316104-132946/> (23/06/04).

[Cámara,02] Cámara, Juan Carlos. Learning Metadata Standards. UPF, Barcelona, 2002. Disponible en: http://www.iaa.upf.es/~jblat/material/doctorat/students/jccbis/Tecnologias_XML.htm (23/06/04).

¹⁰ ARIADNE, Foundation for the European Knowledge Tool. Metadatos. Disponible en: <http://www.ariadne-eu.org/en/publications/metadata/index.html> (4/07/04)

[Dekkers,03] Dekkers, Maks. Dublin core: Metadatos para la interoperabilidad (SeminarioJunio 2003). Disponible en: http://hipatia.uc3m.es/es/eventos/dcmi-es1/Dekkers_spa.pdf (23/06/04).